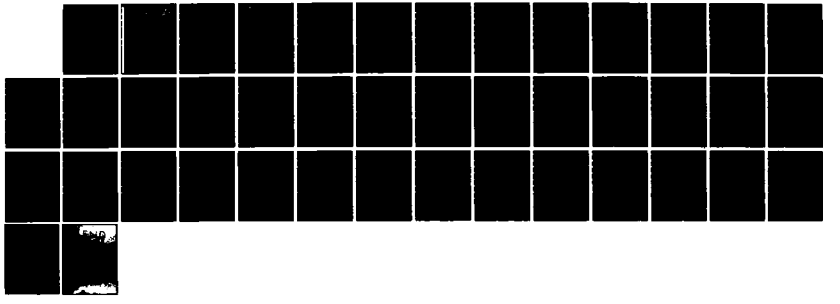
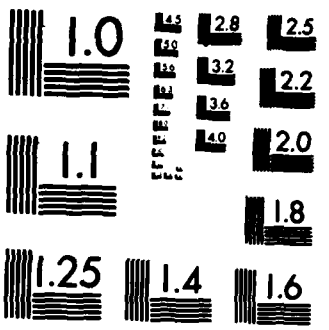
AD-A147 198 BAYESIAN ESTIMATION OF UNDETECTED ERRORS(U) CALIFORNIA UNIV BERKELEY OPERATIONS RESEARCH CENTER W S JEWELL OCT 83 ORC-83-11 AFOSR-81-0122 1/1

UNCLASSIFIED F/G 12/1 NL




1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5
2.8
3.2
3.6
4.0

OPERATIONS RESEARCH CENTER

BAYESIAN ESTIMATION OF UNDETECTED ERRORS

by

William S. Jewell*

ORC 83-11

October 1983

UNIVERSITY OF CALIFORNIA

BERKELEY

BAYESIAN ESTIMATION OF UNDETECTED ERRORS

by

William S. Jewell[†]

ORC 83-11 October 1983

[†]Department of Industrial Engineering and Operations Research, University of California, Berkeley, California.

This research was supported by the Air Force Office of Scientific Research (AFSC), USAF, under Grant AFOSR-81-0122 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ORC 83-11	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) BAYESIAN ESTIMATION OF UNDETECTED ERRORS		5. TYPE OF REPORT & PERIOD COVERED Technical Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) William S. Jewell		8. CONTRACT OR GRANT NUMBER(s) AFOSR-81-0122
9. PERFORMING ORGANIZATION NAME AND ADDRESS Operations Research Center University of California Berkeley, California 94720		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS United States Air Force Air Force Office of Scientific Research Bolling Air Force Base, D.C. 20332		12. REPORT DATE October 1983 13. NUMBER OF PAGES 40
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Bayesian Estimation
Software Reliability Quality Control
Error Detection
Capture-Recapture Sampling
Animal Abundance Studies
Credibility Predictions

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

(SEE ABSTRACT)

ABSTRACT

An unknown number, N, of errors exist in a certain product, for example, defects in a production lot, errors in a manuscript, or bugs in a computer program. I inspectors with possibly different competencies are to be put to work to find the errors. How should the inspection be organized, and what is a good estimate of the undetected errors (or of N)?

This problem is similar to the capture-recapture sampling problem of population biology, assuming a closed population and a parallel search effort, for which many classical results are available. For example, in the case $I = 2$, the Petersen method estimates N as:

$$\hat{N} = \frac{n(1)n(2)}{n_{12}},$$

where $n(i)$ is the total number of errors found by inspector i, $(i = 1,2,)$ and n_{12} is the number of defects found by both inspectors. A corresponding maximum-likelihood estimate of N in the general case is due to Chapman and Darroch, and must be solved recursively (see Seber (1982) for a summary of animal census methods).

Apart from an elementary analysis of the $I = 2$ case by Gaskell and George (1972) and some sequential sampling plans by Freeman (1973), the only Bayesian approach to this problem appears to be by Castledine (1981), who obtains rather complicated results appropriate to the population biology model.

~~In our~~ This paper, we develops the model in a manner more related to error detection problems by first assuming that N is Poisson with parameter λ, and the detection of defects follows a multinomial law, with independent detection probabilities, p_i $(i = 1,2, \ldots, I)$. The maximum likelihood estimator of λ has the same form as the Chapman-Darroch estimator, and a similar result obtains for $Q = \Pi\ (1 - p_i)$, the probability that a given error is overlooked during parallel search.

Next, we analyze the problem in which λ and the p_i are random quantities, by assuming that they are Gamma- and Beta-distributed, respectively. The resulting prediction of the number of unfound errors (the mean of the predictive distribution) can then be expressed as a weighted sum of products of linear "credibility" predictions for λ and the p_i. Surprisingly, the predictive density can be calculated exactly through a recursive relationship which shows that the density is negative binomial in the tails. In the limit, as the prior variances of λ and the p_i increase without bound, the predictive mode approaches the Chapman-Darroch estimator; if we have strong prior information, the mode is given by a generalized Chapman-Darroch form involving credibility formulae.

BAYESIAN ESTIMATION OF UNDETECTED ERRORS

by

William S. Jewell

1. INTRODUCTION

A number of estimation problems in reliability can be described as follows: a certain product has an unknown number, N, of defects. A group of I inspectors each allocates a given amount of independent effort to finding and removing the defects. After finding, say, n_T total defects, what is the estimated number, $n_o = N - n_T$, of *undetected defects* still left in the product?

For example, in manufacturing quality control, the product may be a certain production lot for which the inspectors may use visual or machine-aided techniques to inspect a portion or all of the items. Estimation of the number of undiscovered defects in the sample scrutinized is the first step in setting quality assurance levels for the entire lot.

In software reliability, the defects correspond to program errors or bugs that can be detected and removed by programmers using some combination of visual scanning of program code and of experimental running of the program on typical input. The estimation of undetected errors remaining in the program not only helps certify the application-readiness of the software, but also provides an indication of the effort that will be needed for customer support and for the upgrading of future program releases. A similar interpretation arises in the proofreading of manuscripts for misprints.

Superficially, this model is similar to the problem of estimating the ultimate failure rate of a product during the reliability growth (learning curve) phase of product testing and development (see, e.g., Jewell (1982)).

However, in that application, an unspecified external process of design improvement reduces the stochastic rate of recurrence of product "failures" according to some given law, whose parameters are to be estimated. In this model, on the other hand, an inspector is assumed to actually remove (or at least to identify) one of a finite number of defects or errors, so that, at the end of inspection, there remain only a smaller number of unfound errors. Further, as we shall see below, there is an advantage to having the inspectors work in parallel on the same product, rather than in series, as this helps make more precise any uncertainty in the inspection efficiencies of the different examiners, and thus improves the estimate of undetected errors.

After specifying the basic model, we find first a simple point estimator for N that was originally developed in the field of population biology (by Petersen, Chapman, Darroch, and others) for estimating the size of a closed animal population through capture-recapture sampling. We then make the additional assumption that $\tilde{N}$ is Poisson with parameter λ, and show that the MLE for λ has the same form as the classical estimator of N.

We then analyze the problem from a Bayesian point of view by assuming that $\tilde{\lambda}$ and the detection probabilities for each inspector are random quantities with Gamma- and Beta-prior densities, respectively. After computing the rather complex posterior densities of the parameters, we then find a simpler expression for the predictive density of $\tilde{n}_o$ in recursive form, showing that this density is Negative Binomial in the tails. Moments of this predictive density can only be expressed as a ratio of complex weighted sums of products of linear "credibility" predictors for $\tilde{\lambda}$ and the detection probabilities; however, the posterior mode of $\tilde{n}_o$ can be rearranged into the form of a generalized Petersen-Chapman-Darroch estimator using credibility formulae.

The paper concludes with examples of numerical calculations of the predictive density and remarks on model extensions.

I would like to express my appreciation to Sheldon Ross, who introduced me to this problem area through the paper of Polya (1976), and to Dennis Lindley, who pointed out the connection with capture-recapture census methods.

2. BASIC MODEL; SERIES AND PARALLEL SEARCH STRATEGIES

Suppose that the error inspection process is such that:

(a) Each error present has the same probability of being detected by a given inspector;

(b) The probability that inspector i will find any given error is p_i, $(i = 1,2, \ldots, I)$, independent of previous errors found by i or by any other inspector.

The simplest possible strategy for organizing a search by I inspectors is a *serial* one, in which: inspector #1 examines the raw product (which has an unknown number, N, of errors), and finds and removes n_1 errors; inspector #2 then examines the product (which now has $N - n_1$ errors), finding and removing n_2 errors;.....until the I^{th} inspector finds and removes n_I of the $N - (n_1 + n_2 + \ldots + n_{I-1})$ errors remaining. It follows from the assumptions above that each of the unknown $\tilde{n}_i$ is conditionally Binomially distributed with parameters $(p_i, N - (n_1 + n_2 + \ldots + n_{i-1}))$, so that the joint conditional density of the I pieces of data, i.e., of $(\tilde{n}_1, \tilde{n}_2, \ldots, \tilde{n}_I \mid N ; \underline{p})$, where $\underline{p} = (p_1, p_2, \ldots, p_I)$, is easily found. The total number of detected errors in serial search is $n_T = n_1 + n_2 + \ldots + n_I$, so the number of undetected errors is $n_o = N - n_T$. More importantly, since each error, if present, is missed by i with probability $q_i = 1 - p_i$ $(i = 1,2, \ldots, I)$, the total overlook probability (probability of being undetected by any inspector) for every error is $Q = \prod_{i=1}^{I} q_i$, and thus the conditional density of undetected errors, $(\tilde{n}_o \mid N ; \underline{p})$, is Binomial$(Q,N)$.

A *parallel* search strategy is more complicated, since here we assume, either that the inspectors all work independently on identical copies of the

product, or that they work in some sequence on a single product, (secretly) identifying, but not removing, the defects which they find. With this strategy, there will usually be duplication in the defects found by different inspectors, and the lists of defects reported by each will have to be reconciled, classifying and counting the errors in the following mutually exclusive and collectively exhaustive categories:

n_i - the number of defects found only by inspector i ;

n_{ij} - the number of defects found jointly only by i and j $(i < j)$;

n_{ijk} - the number of defects found jointly only by i, j, and k $(i < j < k)$;

$\vdots \qquad\qquad \vdots \qquad\qquad \vdots$

$n_{123\ldots I}$ - the number of defects found jointly by all inspectors.

Thus, there will be $2^I - 1$ separate pieces of observed data:

$$\mathcal{D} = \{(n_i);(n_{ij});(n_{ijk});\ldots;n_{123\ldots I}\} .$$

Inspector i finds, in total:

$$(2.1) \qquad n(i) = n_i + \sum_j n_{ij} + \sum_{j<k} \sum n_{ijk} + \ldots + n_{123\ldots I}$$

defects, and the total number of distinct defects found by all inspectors is:

$$(2.2) \quad n_T = \sum_i n_i + \sum_{i<j} \sum n_{ij} + \sum_{i<j<k} \sum \sum n_{ijk} + \ldots + n_{123\ldots I}$$

$$= \sum_i n(i) - 1 \sum_{i<j} \sum n_{ij} - 2 \sum_{i<j<k} \sum \sum n_{ijk} - \ldots - (I-1)n_{123\ldots I} .$$

The joint conditional density of $\mathcal{D}$ and $n_o = N - n_T$ is derived in the next section. Note that, in spite of the additional complexity of parallel

search, it again follows from the assumptions that the total overlook probability for each error is Q, and hence $(\tilde{n}_o \mid N ; \underline{p})$ is again Binomial(Q,N). Thus, for fixed N and $\underline{p}$, the density of undetected errors is independent of the search strategy*.

Why, then, would one be interested in parallel search? The answer lies in the fact that, by permitting duplicate errors to be found, we gain additional information about the detection probabilities (p_i), so that if they are unknown quantities at the beginning of inspection, the increased data set associated with parallel search will provide increased precision in the posterior densities of both $\tilde{\underline{p}}$ and $\tilde{n}_o$. Henceforth, we shall assume that a parallel search for errors has been made.

*Of course, if defect removal occupies a substantial portion of the inspection effort, then the two search strategies are no longer comparable in the sense described above.

3. THE PETERSEN-CHAPMAN-DARROCH ESTIMATORS

We begin by deriving some classical point estimators for N using heuristic arguments. For $I = 2$, we can argue as in Polya (1976) that, if p_1 were known, $\hat{N} = n(1)/p_1$ is a reasonable point estimate of the unknown total number of errors. On the other hand, there is also the estimator $\hat{p}_1 = n_{12}/n(2)$ for the first detection probability, since, of the $n(2)$ total errors found by the second inspector, n_{12} were also found by the first. Combining these two estimates, we have:

$$(3.1) \qquad \hat{N} = \frac{n(1)n(2)}{n_{12}} = n_T + \frac{n_1 n_2}{n_{12}} \quad ; \quad \hat{p}_i = \frac{n(i)}{\hat{N}} \qquad (i = 1,2) \ .$$

Note that this argument is symmetric with respect to the two inspectors, and that both singly-found and jointly-found defects are important.

With $I > 2$, a slightly different argument is needed. Let the unknown $\tilde{N}$ be decomposed into found and unfound defects, and replace the latter by its mean value with fixed $\underline{p}$:

$$\tilde{N} = n_T + \tilde{n}_o \simeq n_T + Q\tilde{N} \ .$$

However, the unknown miss probabilities, $\tilde{q}_i$, can be estimated for fixed N by $1 - (n(i)/N)$, so that, combining the two estimates:

$$(3.2) \qquad \hat{N} = n_T + \hat{N} \prod_{i=1}^{I} \left(1 - \frac{n(i)}{\hat{N}}\right) ;$$

$$(3.3) \qquad \hat{p}_i = \frac{n(i)}{\hat{N}} \qquad (i = 1,2, \ldots, I) \quad ; \quad \hat{Q} = 1 - \frac{n_T}{\hat{N}} \ .$$

For $I = 2$, these formulae reduce to (3.1), while for $I = 3$, they require the solution of a quadratic equation, etc. A variety of approximating and iterative procedures are available for (3.2); see Seber (1982). A good initial approximation in the general case is:

$$\hat{N} = \frac{\sum\sum_{i<j} n(i)n(j)}{\sum\sum_{i<j} n_{ij}}, \tag{3.4}$$

which is reminiscent of (3.1). It is easy to show that, if all $n(i)$ are equal to each other and to n_T, then $\hat{N} = n_T$; otherwise, (3.2) has a unique finite root $\hat{N} > n_T$.

In spite of the appearance of n_{ij} in (3.1) and (3.4), it should be clear from (3.2) that only the $I + 1$ pieces of information in the *reduced data set*, $\mathcal{D}^* = \{(n(i)); n_T\}$, are needed to estimate $\tilde{N}$.

(3.1) has a long history in the statistical literature; it was apparently first used by LaPlace in 1783 to estimate the population of France. In population biology, it arises in the capture-recapture sampling of a fixed, but unknown animal population, in which $n(1)$ animals are captured and marked in some distinctive fashion, and then released to mix with the general population. At some later time, when ideal mixing is thought to have occurred, a second sample of $n(2)$ animals are recaptured, of which n_{12} are observed to be already marked. $\hat{N}$ then estimates the total animal census, and is generally called the (C.G.J.) Petersen method, after the Danish fishery biologist who used it to study plaice populations in 1889; however, it is also attributed to a Norwegian, K. Dahl, in 1917, and by ornithologists, to an American, F.C. Lincoln, who calculated waterfowl abundance in 1930. Further details may be found in Seber (1982).

The case $I > 2$ corresponds to a multiple capture-recapture sampling of a closed population, in which successive catches are distinctively (re)marked and then released, in what is called a Schnabel census. The estimator (3.2) was first obtained by Chapman (1952), thus showing that the reduced data set $\mathcal{D}^*$ is sufficient for N and $\underline{p}$; in animal census terminology, this means that the complete capture history, e.g., distinctive remarking, is not needed to estimate the size of a closed population. Darroch (1958) then clarified the derivation of (3.2) and analyzed its properties. Since that time, there has been an explosion of generalizations of this approach in the biometric literature, as well as adaptations to epidemiology and other fields; again, Seber (1982) provides the most convenient summary. By far, the literature uses classical estimation techniques; the Bayesian literature is described below in Section 9.

4. NUMERICAL BEHAVIOR OF THE CLASSICAL ESTIMATOR

To obtain some idea of the empirical properties of $\hat{N}$ (and hence of $\hat{\lambda}$ in (5.5)), simulations of the error detection process were run with a true value of $N = 100$, for $I = 2,4,$ and 8 inspections, and with a range of common detection probabilities, $p_i = p = .05,(.05)\ .30,(.10)\ .90$. 100 samples provided sufficient stability for large p, but 200 samples were needed for smaller values of p, as often the estimator did not exist for small p because no overlap in detection occurred.

The results are summarized in Figures 1,2, and 3. For very small values of p, $\hat{N}$ is badly underbiased, then swings briefly to overbiased values before settling down to the true value as p approaches unity. This effect occurs at lower values of p, and is reduced in magnitude, by increasing I. However, looking at the quantiles, we see that the distribution of possible values of $\hat{N}$ is very unstable, and probably unacceptable, for low values of p.

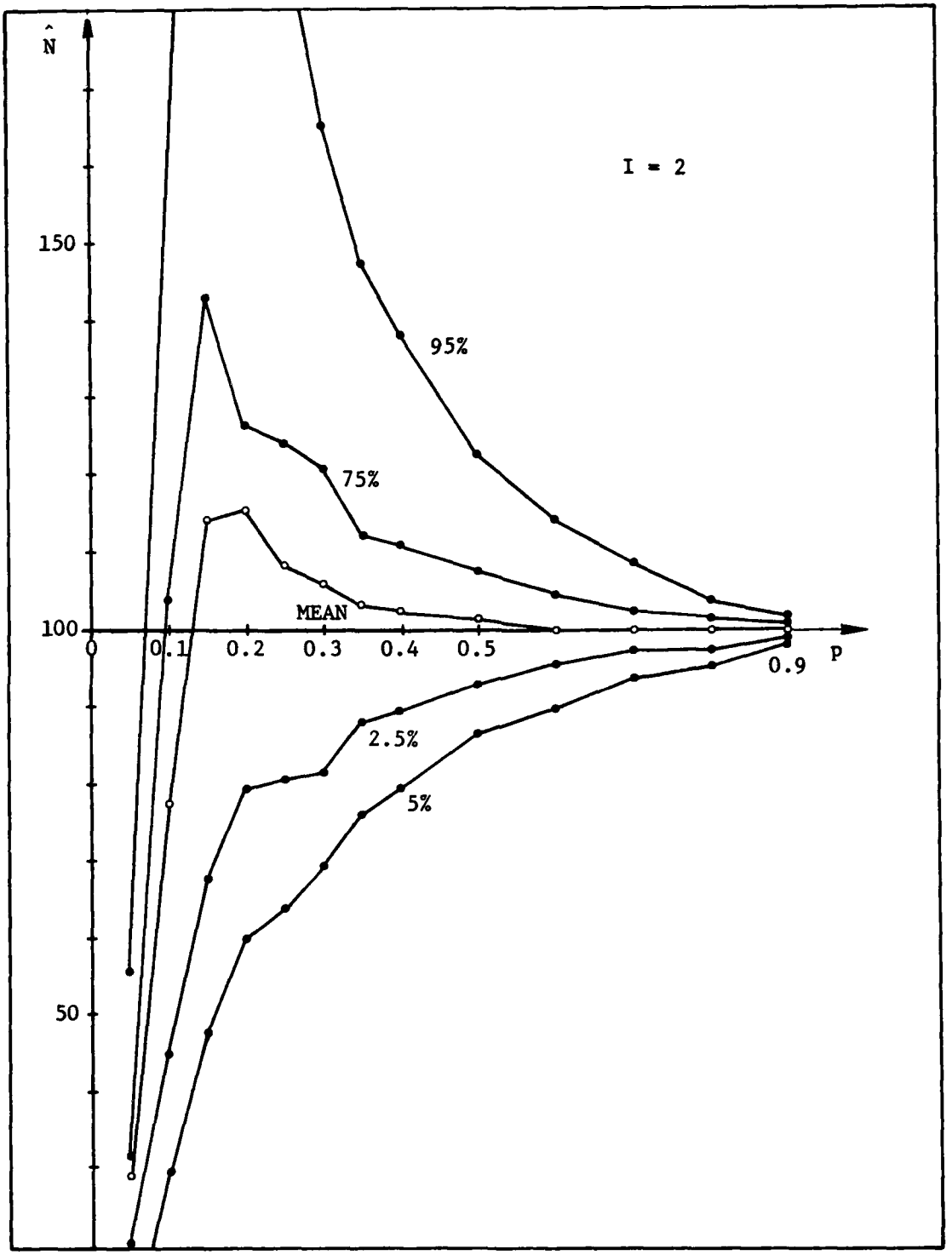


FIGURE 1. Empirical Behavior of Classical Estimator $\hat{N}$ versus Common Detection Probability p, with $I = 2$ Inspectors. $N_{true} = 100$.

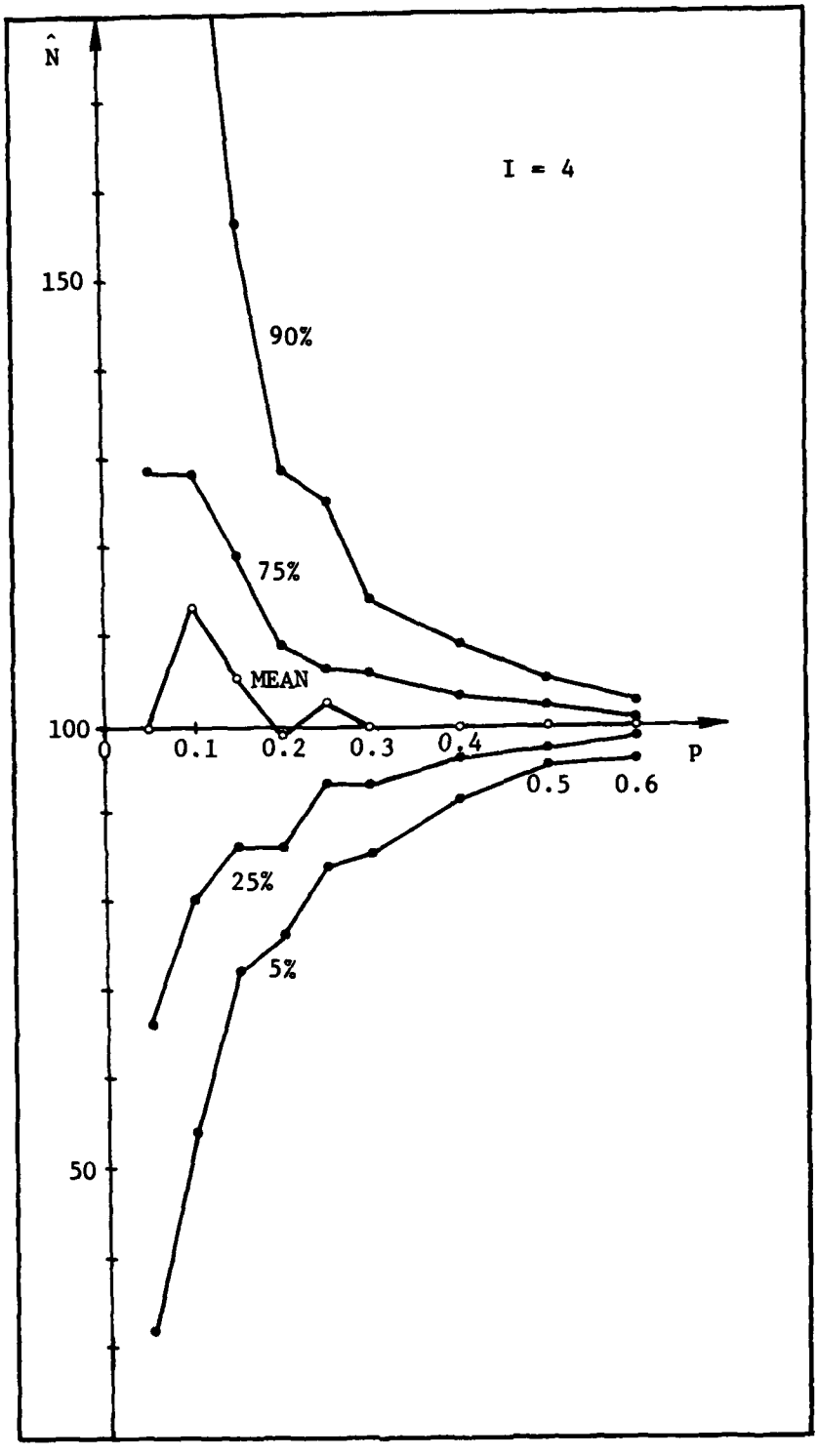


FIGURE 2. Empirical Behavior of Classical Estimator $\hat{N}$ versus Common Detection Probability p, with $I = 4$ Inspectors. $N_{true} = 100$.

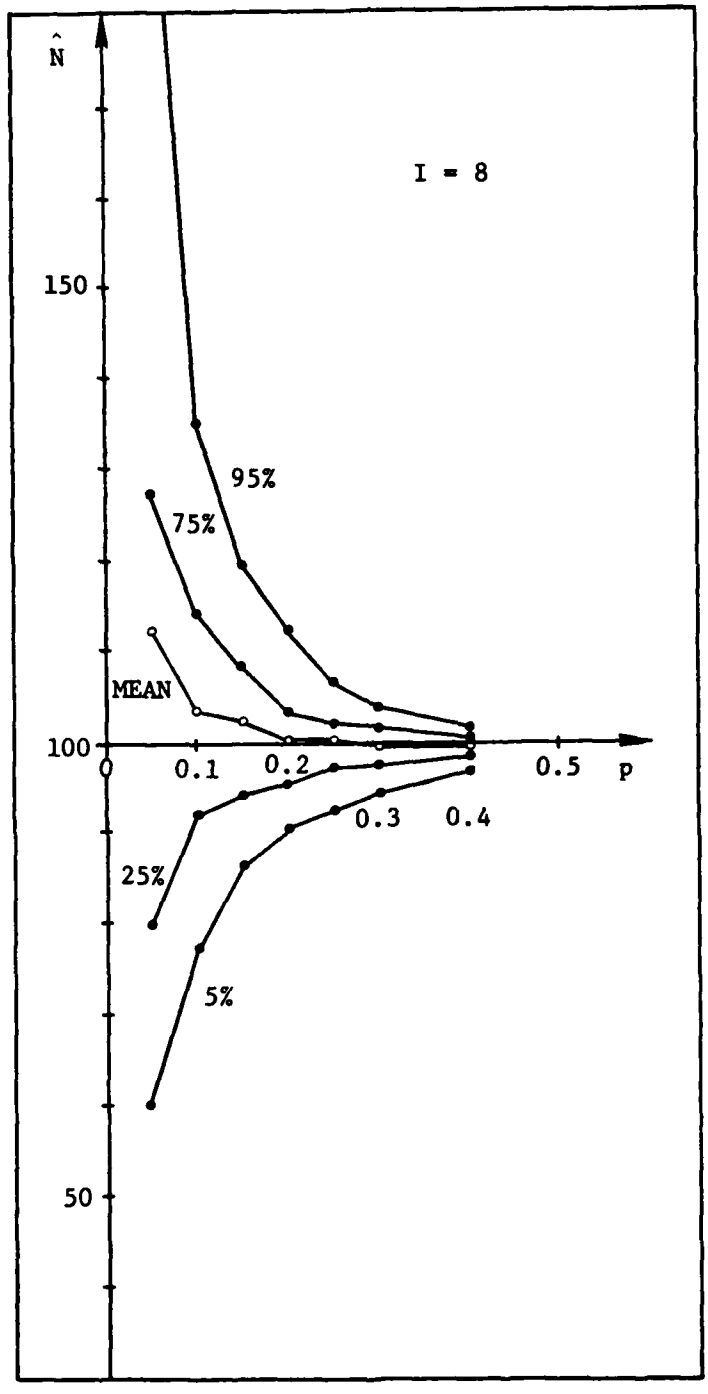


FIGURE 3. Empirical Behavior of Classical Estimator $\hat{N}$ versus Common Detection Probability p, with $I = 8$ Inspectors. $N_{true} = 100$.

5. POISSON ERRORS

Using the assumptions of the last section, it follows that the joint density of $\tilde{n}_o$ and $\tilde{\mathcal{D}}$ under parallel search is the multinomial distribution:

$$p(n_o,\mathcal{D} \mid N , \underline{p}) = \binom{N}{n_o\ \mathcal{D}} Q^{n_o} \prod_{i=1}^{I} (\omega_i Q)^{n_i} , \tag{5.1}$$

where $\omega_i = p_i/q_i$ is the odds-ratio for inspector i, and, of course, $n_o = N - n_T$, with n_T given by (2.2). As might have been expected, (5.1) then simplifies to:

$$p(n_o,\mathcal{D} \mid N , \underline{p}) = \binom{N}{n_o\ \mathcal{D}} Q^{N} \prod_{i=1}^{I} \omega_i^{n(i)} . \tag{5.2}$$

In the reliability applications of interest, it seems natural to assume that the total number of defects or errors would be generated by a Poisson process, with parameter λ ; in the next section, additional modelling flexibility will be added by permitting both λ and $\underline{p}$ to be random quantities. With the Poisson assumption, and eliminating N in favor of n_o, the joint density can be rearranged into:

$$p(n_o,\mathcal{D} \mid \lambda , \underline{p}) \propto \frac{(\lambda Q)^{n_o} e^{-\lambda}}{n_o!} \lambda^{n_T} \prod_{i=1}^{I} p_i^{n(i)} q_i^{n_T - n(i)} .$$

It follows that the conditional density of undetected errors is Poisson(λQ):

$$p(n_o \mid \lambda , p) = \frac{(\lambda Q)^{n_o} e^{-\lambda Q}}{n_o!} , \tag{5.3}$$

(which is expected from first principles), and the data likelihood is:

$$(5.4) \qquad p(\mathcal{D}^* \mid \lambda , \underline{p}) \propto \lambda^{n_T} e^{-\lambda(1-Q)} \prod_{i=1}^{I} p_i^{n(i)} q_i^{n_T - n(i)} \quad .$$

In other words, the reduced data set, $\mathcal{D}^* = \{(n(i)); n_T\}$ is sufficient for both λ and $\underline{p}$.

The maximum-likelihood estimates of the parameters are now:

$$(5.5) \qquad \hat{\lambda} = n_T + \hat{\lambda} \prod_{i=1}^{I} \left(1 - \frac{n(i)}{\hat{\lambda}}\right) ;$$

$$(5.6) \qquad \hat{p}_i = \frac{n(i)}{\hat{\lambda}} \quad (i = 1,2, \ldots, I) \quad ; \quad \hat{Q} = 1 - \frac{n_T}{\hat{\lambda}} ;$$

which can be compared with (3.2)(3.3). In other words, the MLE of λ is exactly the Petersen-Chapman-Darroch estimator for N .

6. A BAYESIAN MODEL

The model just developed is unsatisfactory in most applications because (5.3) depends upon λ and $\underline{p}$ being known exactly, thus giving, for example, $E\{\tilde{n}_o \mid \lambda, \underline{p}\} = \lambda Q$. Usually, these parameters will not be known precisely, and so we will henceforth assume that these are random quantities, with given prior distributions. In this way, the search for errors will also provide us with updated estimates of the rate of error occurrence for this particular product and for the current inspector performance parameters.

Because of the complexity of our final results, even with simple priors, we begin first with cases in which either λ or $\underline{p}$ are known, a priori. This permits us to review known results on appropriate natural conjugate priors, and to suggest methods for estimating hyperparameters. For simplicity, hyperparameters are omitted as explicit arguments, except in priors.

One special notation is convenient in the sequel. If the predictive mean of some random variable $\tilde{y}$ is a linear function of a "natural estimator", $\hat{y} = \hat{y}(\mathcal{D}_y)$ formed from the data, $\mathcal{D}_y$, then we refer to the formula for the predictive mean as a "credibility estimator", because it generally has the form:

$$E\{\tilde{y} \mid \mathcal{D}_y\} = (1 - Z)E\{\tilde{y}\} + Z\hat{y}(\mathcal{D}_y) \stackrel{\text{def}}{=} f_{\tilde{y}}(\hat{y};n,\nu) , \tag{6.1a}$$

where

$$Z = Z(n,\nu) = \frac{n}{n + \nu} \tag{6.1b}$$

is the "credibility factor" which mixes the prior mean, $E\{\tilde{y}\}$, and the natural estimator, $\hat{y}(\mathcal{D}_y)$. n is the "equivalent number of samples" in the data, and ν is the "credibility time constant". This terminology is from

the field of actuarial science, but formulae of this type occur repeatedly in Bayesian prediction or in least-squared approximations to predictive means. A complete bibliography of credibility theory through 1981 is promised as a forthcoming special issue of *Insurance Abstracts and Reviews*.

6.1 Random Error Occurrence Rate, Fixed Detection Probabilities

6.1.1 A Simplified Experiment

Consider first a simplified experiment in which an integer-valued random variable, $\tilde{n}$, is Poisson-distributed, with the mean rate, λ, now considered to be a random quantity. A convenient prior on $\tilde{\lambda}$ is the Gamma(a,b) density:

$$p(\lambda \mid a,b) = \frac{b^a \lambda^{a-1} e^{-b\lambda}}{\Gamma(a)} \qquad (\lambda \geq 0) . \tag{6.2}$$

The hyperparameters (a,b) can be selected by estimating the first two prior moments of the error occurrence rate, since $E\{\tilde{\lambda}\} = a/b$, and $V\{\tilde{\lambda}\} = E\{\tilde{\lambda}\}/b$. (Note that randomizing on $\tilde{\lambda}$ is tantamount to saying that, a priori, $\tilde{n}$ is Negative Binomial$(a, (b+1)^{-1})$.)

(6.2) is advantageous because the Gamma family is "closed under sampling", that is, if the outcome of a single experiment is $\tilde{n} = n_T$, then the posterior-to-data density, $p(\lambda \mid n_T)$ is just Gamma$(a + n_T, b + 1)$, i.e., has the same form as (6.2), but with updated parameters. This simplicity also extends to the posterior-to-data predictive mean of $\tilde{\lambda}$, which is of credibility form:

$$E\{\tilde{\lambda} \mid n_T\} = f_{\tilde{\lambda}}(n_T;1,b) .$$

If a and b are varied, keeping the prior mean, $E\{\tilde{\lambda}\} = a/b$ constant, we see that the time constant, $\nu = b$, shifts the credibility weight, $Z = (1 + b)^{-1}$, to be attached to the outcome of the experiment as a "credible" measure of $\tilde{\lambda}$. (In this sense, the credibility notation hides the fact that $E\{\tilde{\lambda} \mid n_T\}$ depends on *both* a and b.)

6.1.2 Undetected Error Likelihood and Posterior Parameter Density

For our undetected error model, we use the more complex likelihood in (5.4), keeping $\underline{p}$ fixed, which becomes:

$$p(n_T \mid \lambda , \underline{p}) \propto \lambda^{n_T} e^{-\lambda(1-Q)} . \tag{6.3}$$

This modifies slightly the results of the last section, and we find that the posterior-to-data density of $\tilde{\lambda}$ is Gamma$(a + n_T , b + 1 - Q)$. The mean posterior value of the parameter is:

$$E\{\tilde{\lambda} \mid \mathcal{D}^* , \underline{p}\} = E\{\tilde{\lambda} \mid n_T , Q\} = f_{\tilde{\lambda}}(n_T ; (1 - Q) , b) , \tag{6.4}$$

so that, relative to our simplified experiment, the time constant is unchanged, but the equivalent number of samples is reduced. Since Z depends upon the ratio (n/ν), somewhat less credibility is attached to the observation in this model.

6.1.3 Prediction of Undetected Errors

However, the posterior density of $\tilde{\lambda}$ is only an intermediate step to the result of interest, namely, determining the predictive density of $\tilde{n}_o$. Since $p(n_o \mid \lambda , \underline{p})$ is Poisson(λQ) (5.3), the marginal (prior) density of $\tilde{n}_o$, found by mixing with (6.2), is:

$$(6.5)\quad p(n_o \mid Q) = \frac{\Gamma(a + n_o)}{\Gamma(a) n_o!} \left(\frac{b}{b + Q}\right)^a \left(\frac{Q}{b + Q}\right)^{n_o}, \qquad (n_o = 0,1,2, \ldots)$$

that is, Negative Binomial(a , Q/(b + Q)) . Thus, *before* detecting errors, our opinion about the errors that will be undetected *after* the experiment is that $E\{\tilde{n}_o \mid Q\} = aQ/b$, and $V\{\tilde{n}_o \mid Q\} = (aQ/b)\ (1 + (Q/b))$.

From the updating found in Section 6.1.2, it follows that, *after* the detection experiment is over, we will predict that the density of $\tilde{n}_o$, $p(n_o \mid \mathcal{D}^* , \underline{p}) = p(n_o \mid n_T , Q)$ is Negative Binomial with updated parameters $(a + n_T , Q/(b + 1))$. For future reference, this predictive density satisfies the recursion:

$$p(n_o + 1 \mid n_T , Q) = p(n_o \mid n_T , Q)\left(\frac{Q}{n_o + 1}\right)\left(\frac{a + n_T + n_o}{b + 1}\right) .$$

Posterior-to-data, the predicted mean number of defects not yet found is then:

$$(6.6)\quad E\{\tilde{n}_o \mid \mathcal{D}^* , \underline{p}\} = E\{\tilde{n}_o \mid n_T , Q\} = f_{\tilde{n}_o}\left(\left(\frac{Qn_T}{1 - Q}\right) ; (1 - Q) , b\right) .$$

As in (6.4), the hyperparameters enter as the ratio (a/b) in determining $E\{\tilde{n}_o\}$, and b becomes the credibility time constant in a credibility formula with 1 - Q equivalent samples. Thus, if a and b , the parameters of the prior, are large (resp., close to 0), then the natural estimator, $\hat{n}_o = Qn_T/(1 - Q)$, is weakly (resp., strongly) weighted in the prediction, relative to the prior opinion, $E\{\tilde{n}_o\}$. This is exactly what we would expect in comparing the results obtained with strong or weak prior opinion. Finally, note that when $\underline{p}$ is fixed, only Q is in fact used, and only n_T from the data is sufficient for $\tilde{\lambda}$ and $\tilde{n}_o$.

6.2 Random Detection Probabilities, Fixed Error Occurrence Rate

6.2.1 A Simplified Experiment

In the opposite situation, where λ is known, but the $(\tilde{p}_i)$ are jointly random, it is instructive to first consider a simple one-dimensional experiment in which an integer-valued random variable, $\tilde{n} = 0,1,2, \ldots, M$, is Binomial(p,M), with a fixed number of trials, M, but with the success probability, $\tilde{p}$, considered as a random quantity. The convenient natural conjugate prior is the Beta(α,β) density:

$$(6.7) \qquad p(p \mid \alpha,\beta) = B^{-1}(\alpha,\beta)p^{\alpha-1}q^{\beta-1}, \qquad (0 \leq p = 1 - q \leq 1)$$

where B is the Beta function, $B(\alpha,\beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha + \beta)$. Henceforth, we abbreviate $\alpha + \beta = \gamma$. The hyperparameters (α,β) can be selected by estimating the first two moments of the detection probability, since $E\{\tilde{p}\} = 1 - E\{\tilde{q}\} = \alpha/\gamma$, and $V\{\tilde{p}\} = V\{\tilde{q}\} = E\{\tilde{p}\}E\{\tilde{q}\}(\gamma + 1)^{-1}$.

Because (6.7) is closed under sampling relative to the Binomial likelihood, if the outcome of this simplified experiment gives $\tilde{n} = n_S$ successes (and hence $M - n_S$ failures), then the posterior-to-data density of $\tilde{p}$ is again Beta, but with modified hyperparameters $(\alpha + n_S, \beta + M - n_S)$. The posterior-to-data mean predictor of the detection probability is also in credibility form:

$$(6.8) \qquad E\{\tilde{p} \mid n_S\} = 1 - E\{\tilde{q} \mid n_S\} = f_{\tilde{p}}\left(\left(\frac{n_S}{M}\right) ; M , \gamma\right) .$$

The hyperparameter ratio (α/γ) determines the prior mean $E\{\tilde{p}\}$, but in this case it is $\gamma = \alpha + \beta$ which becomes the effective credibility time constant. Naturally, there are M effective trials, with a natural estimator from the data of $\hat{p} = 1 - \hat{q} = n_S/M$.

6.2.2 Joint Detection Probability Prior and Posterior Densities

In our I-dimensional error detection model, it is natural to assume that, a priori, the $(\tilde{p}_i)$ are independently distributed as in (6.7), but with possibly different hyperparameters, so that $p(\underline{p} \mid \underline{\alpha},\underline{\beta})$ is $\prod_{i=1}^{I} \text{Beta}(\alpha_i,\beta_i)$. (With a slight increase in complexity, one could also start with Dirichlet-distributed $\underline{p}$.)

The appropriate part of the data likelihood (5.4), with λ fixed, becomes:

$$(6.9) \qquad p(\mathcal{D}^* \mid \lambda , \underline{p}) \propto e^{-\lambda(1-Q)} \prod_{i=1}^{I} p_i^{n(i)} q_i^{n_T - n(i)} ,$$

and we see that the first term prevents finding a simple natural conjugate prior, and introduces a rather complex coupling between the (p_i) through $Q = \prod q_i$. However, if λ were very small, so that n_T were also small and hence there would be no jointly found errors, then the first term would be approximately unity, and the application of Bayes' law would update each of the hyperparameter groups (α_i,β_i) independently to $(\alpha_i + n(i) , \beta_i + n_T - n(i))$ in a manner similar to the last subsection, with the reasonable interpretation that $n(i)$ is the number of "successes" for inspector i out of n_T total "trials".

For the general case in which λ is of arbitrary size, we expand $e^{+\lambda Q}$ in an infinite series, and find, after some algebra, the (normalized) posterior-to-data joint density of the detection parameters as:

$$(6.10a) \quad p(\underline{p} \mid \mathcal{D}^* , \lambda) = \left[\sum_{k=0}^{\infty} c_k\right]^{-1} \sum_{j=0}^{\infty} c_j \prod_{i=1}^{I} \text{Beta}_i(\alpha_i + n(i) , \beta_i + n_T - n(i) + j) ,$$

with

$$(6.10b)\quad c_j = c_j(\mathcal{D}^*,\lambda) = \frac{\lambda^j}{j!}\prod_{i=1}^{I}\frac{\Gamma(\beta_i + n_T - n(i) + j)}{\Gamma(\beta_i + n_T - n(i))}\frac{\Gamma(\gamma_i + n_T)}{\Gamma(\gamma_i + n_T + j)}\ .$$

Beta_i , of course, refers to the usual one-dimensional Beta density for $\tilde{p}_i$. Marginally, one can find:

$$(6.11)\quad p(p_i \mid \mathcal{D}^*, \lambda) = \left[\sum_{k=0}^{\infty} c_k\right]^{-1} \sum_{j=0}^{\infty} c_j\ \text{Beta}_i(\alpha_i + n(i)\ ,\ \beta_i + n_T - n(i) + j)\ ,$$

but this is misleading, as the $(\tilde{p}_i)$ are now dependent random variables. In general, then, the posterior parameter density is a data-weighted combination of a sequence of simpler experiments, $j = 0,1,2, \ldots,$ in which inspector i has $n(i)$ "successes" out of $n_T + j$ "trials".

The predictive mean of $\tilde{p}_i$ can be expressed as a weighted sum of credibility forms similar to (6.8); for later use, we record the mean predictor for the overall miss probability:

$$(6.12)\quad E\{\tilde{Q} \mid \mathcal{D}^*, \lambda\} = \left[\sum_{k=0}^{\infty} c_k\right]^{-1} \sum_{j=0}^{\infty} c_j \prod_{i=1}^{I} f_{\tilde{q}_i}\left(\left(1 - \frac{n(i)}{(n_T + j)}\right); n_T + j\ ,\ \gamma_i\right)\ ,$$

which has obvious credibility interpretations, in light of the above remarks about "trials" and "successes".

6.2.3 Prediction of Undetected Errors

Since $p(n_o \mid \lambda\ ,\ \underline{p})$ is Poisson(λQ) , one uses the same trick as above to determine the marginal (prior) density of $\tilde{n}_o$:

$$(6.13)\quad p(n_o \mid \lambda) = \left(\frac{\lambda^{n_o}}{n_o!}\right) \sum_{j=1}^{\infty} \frac{(-\lambda)^j}{j!} \prod_{i=1}^{I} \frac{\Gamma(\beta_i + n_o + j)}{\Gamma(\beta_i)} \frac{\Gamma(\gamma_i)}{\Gamma(\gamma_i + n_o + j)} .$$

By previous results, our prior opinion about the mean outcome must be $E\{\tilde{n}_o \mid \lambda\} = \lambda E\{\tilde{Q}\} = \lambda \prod (\beta_i/\gamma_i)$.

Surprisingly, the posterior-to-data predictive density of $\tilde{n}_o$ is simpler than (6.10) or (6.13), as there is a fortuitous cancellation of the term $e^{+\lambda Q}$ in (6.9) with the $e^{-\lambda Q}$ of the Poisson density. After some algebra, we find:

$$(6.14)\quad p(n_o \mid \mathcal{D}^*, \lambda) = p(0 \mid \mathcal{D}^*, \lambda)\left(\frac{\lambda^{n_o}}{n_o!}\right) \prod_{i=1}^{I} \frac{\Gamma(\beta_i + n_T - n(i) + n_o)}{\Gamma(\beta_i + n_T - n(i))} \frac{\Gamma(\gamma_i + n_T)}{\Gamma(\gamma_i + n_T + n_o)} ,$$

which can be put into recursive form, suitable for computation, as:

$$(6.15)\quad \frac{p(n_o + 1 \mid \mathcal{D}^*, \lambda)}{p(n_o \mid \mathcal{D}^*, \lambda)} = \left(\frac{\lambda}{n_o + 1}\right) \prod_{i=1}^{I} f_{\tilde{q}_i}\left(\left(1 - \frac{n(i)}{n_T + n_o}\right) ; n_T + n_o , \gamma_i\right) .$$

(In practice, one sets $p(0 \mid \mathcal{D}^*, \lambda)$ to unity, computes successive probabilities until they become negligible, and then renormalizes. Because of the speed of this method, it appears best to compute the moments of $\tilde{n}_o$ numerically, rather than using, say, $E\{\tilde{n}_o \mid \mathcal{D}^*, \lambda\} = \lambda E\{\tilde{Q} \mid \mathcal{D}^*, \lambda\}$ and (6.12).)

Note the reappearance of the mean credibility predictors for the $(\tilde{q}_i)$, this time with $n(i)$ "successes" out of $n_T + n_o$ "trials" for inspector i . Because the $f_{\tilde{q}_i}$ approach unity as $n_o \to \infty$, the density (6.15) will have a Poisson tail, with parameter λ .

6.3 Random Occurrence Rates and Detection Probabilities

6.3.1 Joint Parameter Prior and Posterior Densities

With these preliminary formulae and interpretations over, we can move quickly through the general case in which both $\tilde{\lambda}$ and the $(\tilde{p}_i)$ are random quantities. For simplicity, we combine the previous priors in an independent manner, so that $p(\lambda\ ;\ \underline{p} \mid a,b\ ;\ \underline{\alpha},\underline{\beta})$ is $\text{Gamma}(a,b) \prod_{i=1}^{I} \text{Beta}_i(\alpha_i,\beta_i)$.

The full form of the likelihood (5.4) must now be used. We note that, *except for the term* $e^{+\lambda Q}$, we would have independent updating of each component of the prior according to:

$$(6.16)\quad a' = a + n_T\ ;\ b' = b + 1\ ;\ \alpha_i' = \alpha_i + n(i)\ ;\ \beta_i' = \beta_i + n_T - n(i)\ ;\ \gamma_i' = \gamma_i + n_T\ .$$

But the coupling term can be expanded into a power series, as in previous subsections, so the posterior-to-data joint parameter density becomes, after normalization:

$$(6.17a)\quad p(\lambda\ ,\ \underline{p} \mid \mathcal{D}^*) = \left[\sum_{k=0}^{\infty} d_k\right]^{-1} \sum_{j=0}^{\infty} d_j\ \text{Gamma}(a'+j\ ,\ b') \prod_{i=1}^{I} \text{Beta}_i(\alpha_i'\ ,\ \beta_i'+j)\ ,$$

with

$$(6.17b)\quad d_j = d_j(\mathcal{D}^*) = \frac{(b')^{-j}}{j!}\ \frac{\Gamma(a'+j)}{\Gamma(a')} \prod_{i=1}^{I} \frac{\Gamma(\beta_i'+j)}{\Gamma(\beta_i')}\ \frac{\Gamma(\gamma_i')}{\Gamma(\gamma_i'+j)}\ .$$

In general, both $\tilde{\lambda}$ and the $(\tilde{p}_i)$ are correlated, a posteriori. Moments of the parameters can now be obtained as in (6.12), provided one can compute the coefficients $(d_j(\mathcal{D}^*))$ (see below).

6.3.2 Prediction of Undetected Errors

Using (5.3), one finds the marginal density of $\tilde{n}_o$, prior to inspecting for errors, to be a rather complex combination of (6.5) and (6.13):

$$(6.18)\quad p(n_o) = \frac{\Gamma(a+n_o)}{\Gamma(a)n_o!}\, b^{-n_o} \sum_{j=0}^{\infty} \frac{(-b)^{-j}}{j!} \frac{\Gamma(a+n_o+j)}{\Gamma(a)} \prod_{i=1}^{I} \frac{\Gamma(\beta_i+n_o+j)}{\Gamma(\beta_i)} \frac{\Gamma(\gamma_i)}{\Gamma(\beta_i+n_o+j)} .$$

Of course, a priori $E\{\tilde{n}_o\} = E\{\tilde{\lambda}\tilde{Q}\} = E\{\tilde{\lambda}\}E\{\tilde{Q}\} = (a/b) \prod (\beta_i/\gamma_i)$.

Again, we are surprised to find that the complexity of (6.17) and (6.18) are not carried over into the predictive density of $\tilde{n}_o$, because of the fortuitous cancellation of two exponential terms. After some algebra, we find:

$$(6.19)\quad p(n_o \mid \mathcal{D}^*) = p(0 \mid \mathcal{D}^*)\left(\frac{(b')^{-n_o}}{n_o!} \frac{\Gamma(a'+n_o)}{\Gamma(a')}\right) \prod_{i=1}^{I} \frac{\Gamma(\beta_i'+n_o)}{\Gamma(\beta_i')} \frac{\Gamma(\gamma_i')}{\Gamma(\gamma_i'+n_o)} ,$$

which should be compared with (6.14) and the updated version of (6.5). In fact, from (6.17), we see that:

$$p(n_o \mid \mathcal{D}^*) = \frac{d_{n_o}(\mathcal{D}^*)}{\sum_{j=0}^{\infty} d_j(\mathcal{D}^*)} ,$$

which gives a logical interpretation to the weights in that formula. Similarly, $E\{\tilde{n}_o \mid \mathcal{D}^*\}$ could also be expressed as the ratio of two weighted sums of products of linear credibility formulae.

As in (6.15), the predictive density can be put into recurrence form as:

$$(6.20)\quad \frac{p(n_o+1 \mid \mathcal{D}^*)}{p(n_o \mid \mathcal{D}^*)} = \left(\frac{1}{n_o+1}\right)\left(\frac{a'+n_o}{b'}\right) \prod_{i=1}^{I} \left(\frac{\beta_i'+n_o}{\gamma_i'+n_o}\right)$$

$$= \left(\frac{1}{n_o+1}\right) f_{\tilde{\lambda}}(n_T+n_o\,;\,1\,,\,b) \prod_{i=1}^{I} f_{\tilde{q}_i}\left(\left(1-\frac{n(i)}{n_T+n_o}\right)\,;\,n_T+n_o\,;\,\gamma_i\right) .$$

Numerical computation is very efficient; by setting $p(0 \mid \mathcal{D}^*) = 1$, one in fact computes the coefficients $d_{n_o}(\mathcal{D}^*)$, and then gets the predictive density through normalization. Moments of $\tilde{n}_o$ are thus best found numerically. As $n_o \to \infty$, $f_{\tilde{q}_i} \to 1$ for every i, so that $p(n_o \mid \mathcal{D}^*)$ has a Negative Binomial$\left(a + n_T\,,\,(b+1)^{-1}\right)$ tail, similar to the predictive density with $\underline{p}$ fixed in Subsection 6.1.3.

Of course, the various special results of Subsections 6.1 and 6.2 can now be gotten from the formulae above through appropriate limiting values of the hyperparameters.

7. THE POSTERIOR MODE

Although (6.20) permits the calculation of any moment of $\tilde{n}_o$, it is difficult to compare these Bayesian results with the classical point estimators of Section 3.

However, the *mode* of the predictive density is easily found as the smallest integer, $\ddot{n}_o$, for which $p(\ddot{n}_o + 1 \mid \mathcal{D}^*) \leq p(n_o \mid \mathcal{D}^*)$. From (6.20), after some rearranging, we find that $\ddot{n}_o$ is the smallest integer not less than the solution n_o^* to:

$$(7.1) \quad n_o^* + 1 = \left\lfloor n_o^* + n_T + \left(\frac{b}{b+1}\right)\left(E\{\tilde{\lambda}\} - n_o^* - n_T\right)\right\rfloor \prod_{i=1}^{I} \left[1 - \frac{n(i) + \gamma_i E\{p_i\}}{n_o^* + n_T + \gamma_i}\right],$$

which should be compared with (3.2), rewritten with $\hat{n}_o = \hat{N} - n_T$:

$$(7.2) \qquad \hat{n}_o = \left\lfloor \hat{n}_o + n_T \right\rfloor \prod_{i=1}^{I} \left[1 - \frac{n(i)}{\hat{n}_o + n_T}\right]$$

whence we can easily see the effect of adding prior opinion.

If $b \to 0$ and $\gamma_i \to 0$, with constant prior means $E\{\tilde{\lambda}\}$ and $E\{p_i\}$, n_o^* approaches $\hat{n}_o$, so that this would correspond to "diffuse" prior knowledge (although, for the Beta density, $\alpha = \beta = 1$ and $\gamma = 2$ is usually considered the diffuse case). Conversely, as $b \to \infty$, with constant $E\{\tilde{\lambda}\}$, the mode approaches the integer above the solution to:

$$(7.3) \qquad n_o^* + 1 = E\{\tilde{\lambda}\} \prod_{i=1}^{I} f_{\tilde{q}_i}\left(\left(1 - \frac{n(i)}{n_o^* + n_T}\right) ; n_o^* + n_T ; \gamma_i\right),$$

or, if all the $\gamma_i \to \infty$, with $E\{p_i\}$ fixed, the mode is the integer above the solution to:

$$n_o^* + 1 = \left(\frac{a + n_T + n_o^*}{b + 1}\right) E\{\tilde{Q}\} \quad , \tag{7.4}$$

which is practically the posterior mean (6.6). Thus, the posterior-to-data mode of the predictive density is intimately related to, and a natural generalization of, the Petersen-Chapman-Darroch estimators.

8. OTHER MODELS

A variety of related error-detection models can be developed using the above methods. For example, in animal census studies, it is often assumed that the capture probability remains constant at each trial; this is the same as assuming that the individual detection probabilities $(\tilde{p}_i)$ are equal to some *common* unknown value, $\tilde{p}$. For this case, one can easily show that the two statistics, $\mathcal{D}^{**} = \left\{ n_T , \bar{n} = I^{-1} \sum n(i) \right\}$, are sufficient, giving estimators:

$$(8.1) \qquad \hat{N} = n_T + \hat{N}\left(1 - \frac{\bar{n}}{\hat{N}}\right)^I \quad ; \quad \hat{q} = 1 - \frac{\bar{n}}{\hat{N}} \quad ; \quad \hat{Q} = \hat{q}^I .$$

If we assume a Beta(α,β) prior for $\tilde{p}$, $\tilde{\lambda}$ remaining Gamma(a,b) a priori, we find that, corresponding to (6.20), the predictive density for undetected errors satisfies the recursion:

$$(8.2) \qquad \frac{p(n_o + 1 \mid \mathcal{D}^{**})}{p(n_o \mid \mathcal{D}^{**})} = \left(\frac{a + n_T + n_o}{b + 1}\right) \prod_{j=0}^{I-1} \left(\frac{\beta + I(n_T - \bar{n} + n_o) + j}{\gamma + I(n_T + n_o) + j}\right) .$$

The convergence of the estimate $\hat{Q}$ with increasing I is quite rapid because of the increased rate of learning about $\tilde{p}$ in this model.

A related variation occurs when the inspectors have a common "unit" detection probability, but expend different known amounts of effort or search duration (e_i); this is tantamount to assuming $\tilde{p}_i = e_i \tilde{p}$ $(i = 1,2, \ldots, I)$.

One can also assume that the error detection or correction process is defective, or that new errors can enter randomly during inspection; this leads to likelihoods related to those already analyzed for non-closed animal population studies (Seber (1982)). Or, one can assume that detection probabilities are different for different error types (Otis et al., (1978)). And so forth.

Finally, one could also make a Bayesian analysis of the serial inspection strategy; however, as explained earlier, we expect this to be less efficient at predicting unfound errors because less information about the unknown detection probabilities is generated. A comparison between these two approaches will be the subject of a forthcoming paper.

9. OTHER BAYESIAN MODELS

Apart from an elementary model for $I = 2$ by Gaskell and George (1972), a partly Bayesian approach by Carle and Strub (1978), a sequential sampling plan for $I = 1$ by Yang et al., (1982), and a sequential sampling plan for $I = 2$ by Freeman (1973), the only general Bayesian model of which the author is aware is by Casteldine (1981). Starting with the likelihood (3.2), he assumes:

$$\text{either that} \quad \text{I: all } \tilde{p}_i = \tilde{p} \text{ , which is Beta}(\alpha,\beta) \text{ ,}$$
$$\text{or that} \quad \text{II: each i.i.d. } \tilde{p}_i \text{ is Beta}(\alpha,\beta) \text{ ,}$$

and that $\tilde{N}$ has an arbitrary independent prior, $\Pi(N)$. From this point on, his argument is mostly numerical or approximative in nature, concentrating on $\Pi(N)$ constant or $\Pi(N) \propto N^{-1}$. Some other more complex variations are also explored, for example, a two-stage model in which $(\ln \tilde{\omega}_i)$ is Normal$(\tilde{\theta},\sigma^2)$, σ^2 known, and $\tilde{\theta}$ is also normally distributed with known hyperparameters. However, additional approximations appear necessary to interpret these variations.

This is in contrast to our results, which require $\tilde{N} \sim$ Poisson$(\tilde{\lambda})$ and $\tilde{\lambda} \sim$ Gamma(a,b) , which is tantamount to assuming $\tilde{N}$ is Negative Binomial, a priori. While this assumption may be of limited validity in animal population studies, it seems like a useful starting point for reliability modelling, at least until empirical error and defects distributions are available (Yang et al., (1982) argue a Gamma-Poisson assumption in proofreading manuscripts). Our predictive densities also have the advantage that they can be expressed in closed form, with "credibility" interpretations for many of the components, and the posterior mode can be related to the classical Petersen-Chapman-Darroch formula.

Other Bayesian variations will, no doubt, also prove useful in application.

10. NUMERICAL BEHAVIOR OF THE BAYESIAN ESTIMATOR

To obtain some idea of the numerical properties of (6.19), simulations were run using various priors, and various values of I .

For the detection probabilities, it was assumed that for the Beta priors, $\alpha_i = \beta_i = 1.0$, which gives uniform densities for all i . Three cases of error rate prior were examined:

	$E\{\tilde{\lambda}\}$	$V\{\tilde{\lambda}\}$
I	50	1250
II	100	5000
III	200	20000

The shape parameter a of the Gamma prior was kept constant at $a = 2$, with b adjusted to give the above moments. Since N_{true} was 100, it can be seen that these correspond to low, O.K., and high prior estimates, though of course $N = 100$ could have occurred from any prior.

Then, one sample of data was obtained for $I = 1,2,4,$ and 8, with assumed values $p_i = 0.5$ for all i . The data sets obtained were:

$I = 1 \quad n_T = 45 \quad \underline{n} = (45)$
$I = 2 \quad n_T = 79 \quad \underline{n} = (55,47)$
$I = 4 \quad n_T = 95 \quad \underline{n} = (48,52,57,45)$
$I = 8 \quad n_T = 99 \quad \underline{n} = (50,55,42,47,50,44,51,50)$.

Of course, the results would have been quite different in another simulation. The classical estimator, $\hat{N}$, does not exist for $I = 1$; but would have given values of 112.39, 101.31, and 99.45, that is, $\hat{n}_o = 33.39$, 6.31, and 0.45 for $I = 2,4,8$ respectively.

Figure 4 shows the density $p(n_o \mid \mathcal{D})$ for $I = 1$, for the three priors given above; the effect of the priors on the predictive mean, though not on

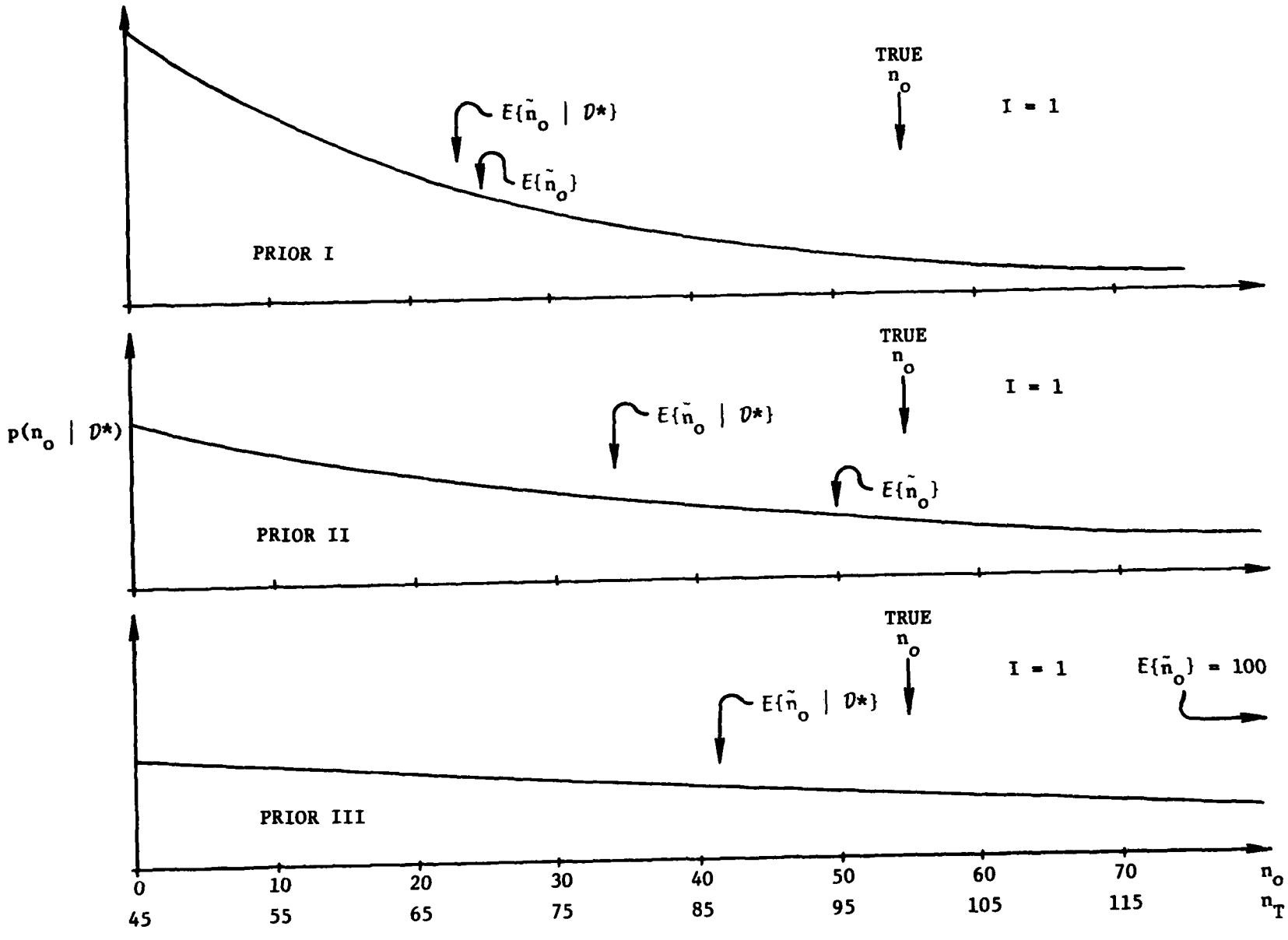


FIGURE 4. Predictive Density for one sample from one observer, three different priors. (Continuous curve approximates discrete density)

the shape can be clearly seen. Figure 5 shows that the predictive density develops an interior mode when $I = 2$, although the difference due to different priors is less perceptible. For $I = 4$ and 8, the effect of the priors is barely perceptible, so that Figure 6 shows just case II above; for $I = 8$, the mode is again at $n_o = 0$.

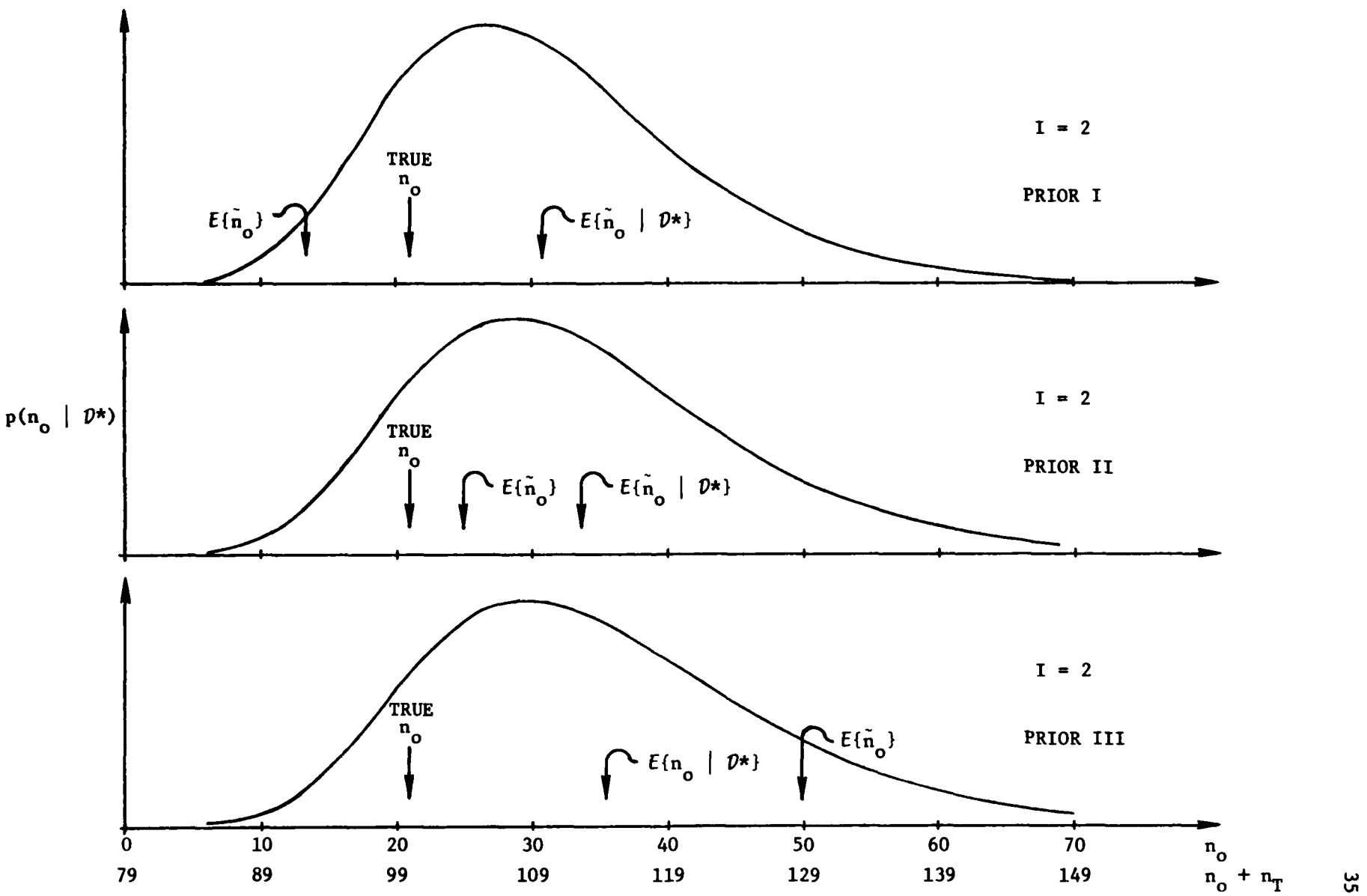


FIGURE 5. Predictive Density for one sample from two observers, three different priors. (Continuous curve approximates discrete density)

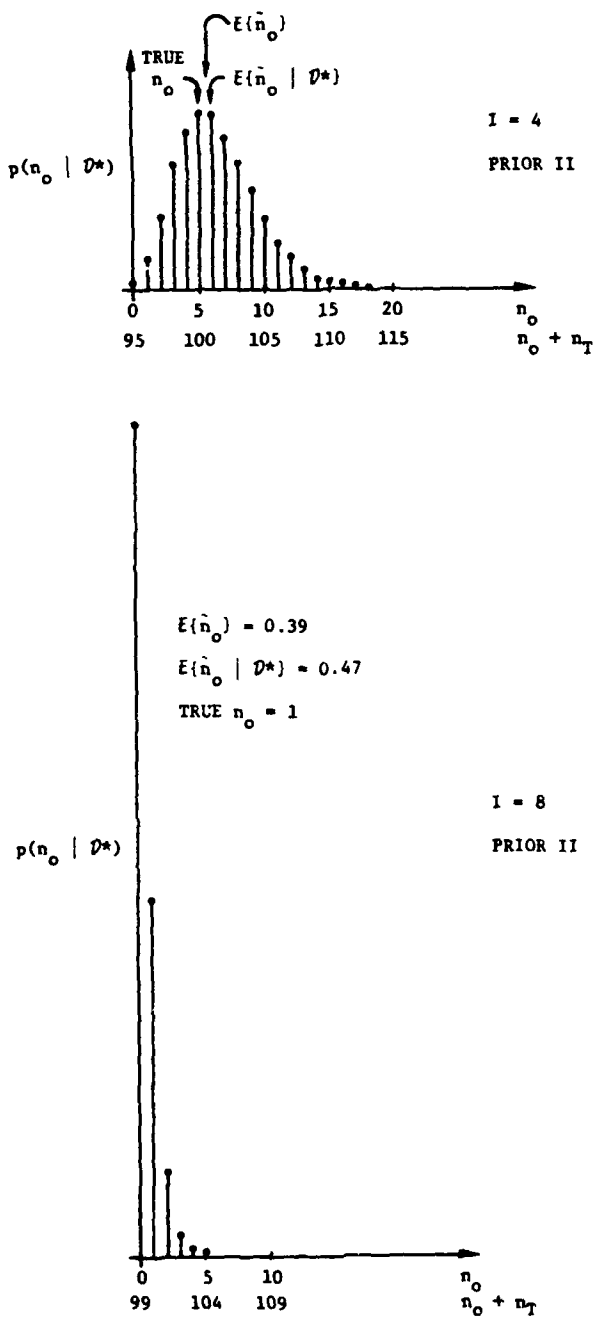


FIGURE 6. Predictive Densities for one sample from four and eight observers, respectively, using prior density II.

REFERENCES

1. Carle, F. L. and M. R. Strub, (1978). "A New Method for Estimating Population Size From Removal Data," Biometrics, 34, 621-30.

2. Chapman, D. G., (1952). "Inverse Multiple and Sequential Sample Censuses," Biometrics, 8, 286-306.

3. Darroch, J. N., (1958). "The Multiple-Recapture Census, I: Estimation of a Closed Population," Biometrika, 45, 343-59.

4. Freeman, P. R., (1973). "Sequential Recapture," Biometrika, 60, 141-53.

5. Gaskell, T. J. and B. J. George, (1972). "A Bayesian Modification of the Lincoln Index," J. Appl. Ecol., 41, 377-84.

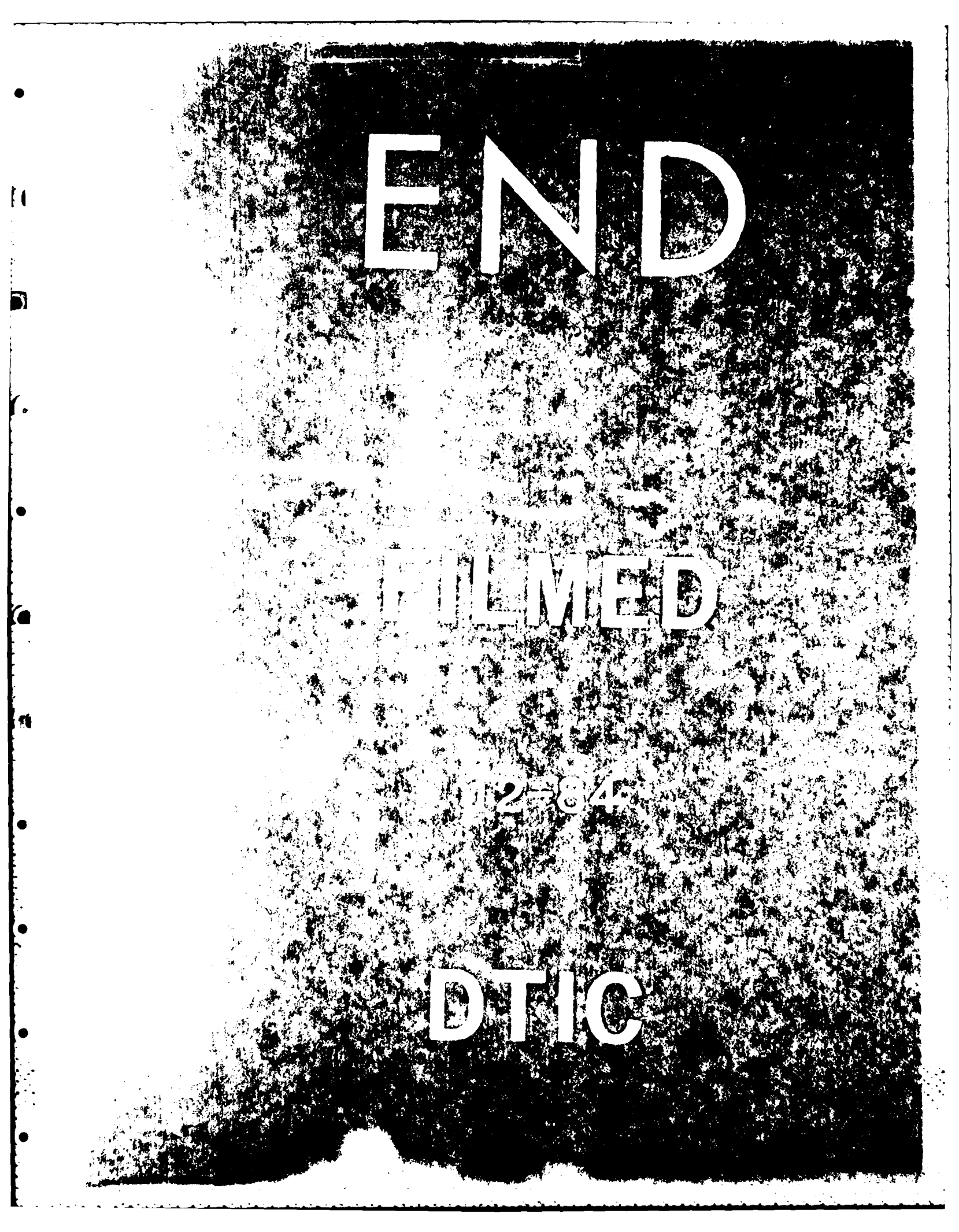
6. Jewell, W. S., (1982). "A General Framework for Learning-Curve Reliability Growth Models," ORC 79-11, Operations Research Center, University of California, Berkeley. To appear, Operations Research, (1984).

7. Otis, D. L., K. P. Burnham, G. C. White and D. R. Anderson, (1978). "Statistical Inference for Capture Data From Closed Populations," Wildlife Monograph No. 62, The Wildlife Society, Washington, D.C.

8. Polya, G., (1976). "Probabilities in Proof [illegible]ing," Amer. Math. Monthly, 83, 42.

9. Seber, G. A. F., (1982). The Estimation of Animal Abundance (Second Edition), MacMillan, New York.

10. Yang, M. C. K., D. D. Wackerly and A. Rosalsky, (1982). "Optimal Stopping Rules in Proofreading," J. Appl. Prob., 19, 723-9.


END
FILMED
2-84
DTIC